SPECIAL ARTICLE

Relationship problems and the DSM: needed improvements and suggested solutions

RICHARD E. HEYMAN[1], AMY M. SMITH SLEP[1], STEVEN R.H. BEACH[2], MARIANNE Z. WAMBOLDT[3], NADINE J. KASLOW[4], DAVID REISS[5]

[1]Family Translational Research Group, Department of Psychology, Stony Brook University, State University of New York, Stony Brook, NY 11794-2500, USA
[2]Institute for Behavioral Research, University of Georgia, Athens, GA 30602, USA
[3]Department of Psychiatry, University of Colorado at Denver, and Health Sciences Center, The Children's Hospital, Denver, CO 80218-1088, USA
[4]Department of Psychiatry and Behavioral Sciences, Grady Health System, Emory University School of Medicine, 80 Jesse Hill Jr. Drive, Atlanta, GA 30303, USA
[5]Child Study Center, Yale University, New Haven, CT 06520-7900, USA

Relational problems are clinically significant behavioral or psychological syndromes or patterns that occur between or among individuals and that are associated with present distress or disability or with a significant increased risk of suffering death, pain, disability, or an important loss of freedom. Relational problems (e.g., partner relational problems, partner abuse, child maltreatment) are included as Axis I disorders in the DSM-IV as V-codes (i.e., "Other conditions that may be a focus of clinical attention"). However, there are no criteria provided in the DSM-IV for these codes. In this article, we briefly review literature that incontrovertibly documents both relational problems' syndromes/patterns and their serious sequelae. We then review a series of studies that provide evidence of content validity and inter-rater agreement for criteria to determine presence versus absence of relational problems. The most studied subset of relational problem criteria, those for partner and child maltreatment, have been shown to have remarkably high levels of reliability when disseminated broadly in the field (κ = .66-.89), at agreement levels never reached by DSM diagnoses for individuals. We conclude by arguing that science, service, families, individuals, and the DSM itself, would be well served to include diagnostic criteria for relational problems and to consider the various options for placement of relational problems/processes in the DSM-V.

Key words: Relational problems, partner and child maltreatment, diagnostic criteria, DSM-IV

(World Psychiatry 2009;8:7-14)

The DSM-IV-TR limits mental disorders, by definition, to problems within a person: "a clinically significant behavioral or psychological syndrome or pattern that occurs *in an individual* and that is associated with present distress (e.g., a painful symptom) or disability (i.e., impairment in one or more important areas of functioning) or with a significantly increased risk of suffering death, pain, disability, or an important loss of freedom" (1).

Clinically significant behavioral or psychological syndromes or patterns that occur *between or among individuals* and that are associated with present distress or disability or with a significant increased risk of suffering death, pain, disability, or an important loss of freedom, receive little attention in the DSM-IV. This article focuses on eleven such syndromes/patterns occurring within families – partner relational problems, parenting problems, parent-child relational problems, partner maltreatment (physical, emotional, and sexual abuse; neglect) and child maltreatment (physical, emotional, and sexual abuse; neglect) – which we will refer to as "relational problems".

We briefly review how relational problems are handled in DSM-IV. We sketch the scientific body of research that supports the importance of relational processes to individual's functioning and well-being. We provide several examples of the way criteria could be presented in a clear, structured manner. We present the series of studies that provide evidence of content validity and inter-rater agreement for criteria to determine presence versus absence of relational problems. We discuss the development of screening and diagnostic interviews for relational problems. We present recommendations for possible inclusion of relational problems/processes in the DSM-V. We conclude that criteria are available that could enhance the description of key relationship problems relevant for the provision of optimal clinical care.

HOW RELATIONAL PROBLEMS ARE HANDLED IN THE DSM-IV

The DSM-IV-TR includes relational processes in the section "Other conditions that may be a focus of clinical attention" (e.g., partner relational problem, parent-child relational problem, problems related to abuse or neglect). Further, it lists categories of psychosocial problems on Axis IV (e.g., problems with primary support group, problems related to social environment) and provides the Global Assessment of Functioning (GAF) scale on Axis V and the Global Assessment of Relational Functioning (GARF) scale in Appendix B. In addition, some relational problems have been addressed in supplemental materials, such as the discussion of abuse and neglect and other relational problems in Volume 3 of the DSM-IV Sourcebook (2). However, there are no criteria provided in the DSM-IV for relational problems (or any V-codes).

The prime reason for including criteria for relational problems in the DSM is that clinicians and researchers routinely assess and treat individuals, couples, and families with relational problems – or individual disorders related to, or exacerbated by, relational problems – but are not provided with any guidance regarding criteria. Studies using varying opera-

tionalizations of relational problems are difficult to compare (3,4); agreement among clinicians in the field is poor (5); and the content validity of typical operationalizations is debated (sometimes quite bitterly in the case of partner maltreatment) (6,7). The result is less than optimal research communication, less than optimal accumulation of research results, and less than optimal clinical practice.

The neo-Kraepelinian approach used to develop the DSM described potentially clinically significant syndromes, with an initial goal of classifying disorders reliably. The validity of such classifications then could be evaluated. As noted by Kupfer et al (8), "from the outset... it was recognized that the primary strength of a descriptive approach was its ability to improve communication among clinicians and researchers, not its established validity". In other words, DSM definitions were bootstrapped, because criteria for clinical syndromes had to precede research on prevalence, etiology, and treatment efficacy and effectiveness (9). With reliable, operationalized syndromes as a starting point, validity research is then possible.

Thus, the real question is not "should relational problems be included in the DSM-V?" but "should there be operationalized criteria?" and "should relational problems remain as V-codes or should they be placed somewhere else?" We will argue strenuously that the answer to the first question is emphatically "yes". We are agnostic on the second question; a detailed consideration of the options for placement of relational problems and relational processes in the DSM can be found elsewhere (10,11).

THE NEED FOR RELIABLE AND VALID CRITERIA FOR RELATIONAL PROBLEMS – CASE STUDY: PARTNER RELATIONAL PROBLEMS

Space constraints preclude even a cursory review of the expansive literature linking each of the eleven relational problems with significant distress, morbidity, and mortality. We have chosen, instead, to briefly present the evidence for one problem: partner relational problems. Equivalent sets of literature could be marshaled for nearly all of the other ten problems.

Partner relational problems' effects on adult mental health

The literature linking adult intimate relationships to mental health outcomes is substantial. There are documented connections between relational processes and the etiology, maintenance, relapse, and optimal treatment of many disorders. Because we do not have the space to review this literature exhaustively, we focus on briefly sampling this literature for illustrative purposes.

Serious partner relationship dissatisfaction predicts increased risk for a major depressive episode in the subsequent year, even after controlling for history of depression (12) or comorbidity (13). Both relationship conflict and physical abuse predict subsequent increases in depressive symptoms among women (14). The effect of humiliating relationship events on depression is substantial (15,16). From a behavioral-genetic perspective, the effect of partner relationship satisfaction is a nonshared environmental effect and is not well modeled as resulting from the same genetic factors that produce the vulnerability for depressive symptoms (17). Accordingly, disturbance in intimate adult relationships is key for understanding the etiology of depressive symptoms for many individuals and has the potential to supplement genetically based models (18).

Treatment approaches targeting intimate relationships have proved useful for, among other individual disorders, depression (19), alcohol abuse (20), and drug abuse (20). There are notable applications of relational interventions for individuals with severe mental illness (21). Such treatments are associated with reduced interpersonal stress, greater medication adherence, and lower rates of rehospitalization. As a result of such links, attention to relational problems has increased in the treatment of many mental health problems and is essential for the appropriate management of a number of disorders.

Partner relational problems' effects on children's mental health

Partner relationship conflict is associated with worse parenting and child adjustment, problematic attachments, and increased parent-child and sibling conflicts. Aspects of relationship conflict that have a particularly negative influence on children include frequent, intense, physical, unresolved, child-related conflicts and conflicts attributed to the child's behavior (22). Relationship and parenting problems can be mutually exacerbating and may work synergistically to create a coercive family environment. In turn, relationship and parenting problems can interact with genetic liabilities and influence gene expression to affect the etiology of many physical and mental disorders. For example, women who were adopted soon after birth and who are at high genetic risk for depression show no evidence of the disorder if reared by adoptive parents without psychopathology or relationship difficulties (23). Similarly, adoptees with a genetic risk for schizophrenia and exposure to specific communication styles in their adopted families are more likely to develop the disorder than genetically susceptible persons raised by families with more clear communication and clear roles (24). These data suggest that an interaction between the adult partner relationship environment and particular genetic diatheses may be critical to the etiology of certain major mental disorders (18).

Animal data also indicate the significance of early rearing environment; for example, poor maternal care by rat dams of their pups within the first 10 days of life influences gene expression (25). Poor maternal care leads to changes in glu-

cocorticoid receptor messenger RNA expression in the hippocampus, resulting in enhanced glucocorticoid feedback sensitivity and increased sensitivity to stress. Such changes are the basis for lifetime sensitivity to stress of the maltreated pups (25) and set the stage for the offspring's own poor maternal care of their young. Conversely, good maternal care of infant monkeys at risk for anxiety symptoms moderates symptom expression (26), suggesting that gene-family environment interactions may transform genetic liabilities into genetic assets and that disturbances in primary relationships early in life can change neural systems that control long-term emotional resilience or vulnerability (27).

EXAMPLES OF CRITERIA FOR RELATIONAL PROBLEMS

The development and validation of criteria for relational problems in the DSM-V is well advanced. The goal has been to produce clear criteria that could provide clinically useful guidance and create a basis for inter-rater agreement in clinical settings. In addition, each criteria set is based on the best available scientific understanding of the development and maintenance of these problems.

Table 1 displays the criteria for partner relational problem.

Table 1 Diagnostic criteria for partner relational problem

A. Relationship dissatisfaction during the past month, as evidenced by any of the following:
1) Pervasive sense of unhappiness with the relationship, more days than not.
2) Thoughts of divorce/separation that are more than transitory.
3) Perceived need for professional help for the relationship.

B. Significant impact of the relational dissatisfaction on behavioral, cognitive, or affective systems, as evidenced by at least one of the following for at least one of the partners:
1) Behavioral symptoms:
   a. Conflict resolution difficulties, as evidenced by either:
      i. Persistent and marked escalation of negative behavior or affect (e.g., "little" disputes quickly and frequently evolve into heated arguments).
      ii. Pervasive withdrawal so that resolution is impeded.

      (Note: Withdrawal can be either through leaving a discussion before it is resolved or through more pervasive disconnectedness that impedes bringing up or resolving problems)
   b. Pervasive lack of positive behaviors (e.g., sharing thoughts and feelings; affection) or supportive behaviors.
2) Cognitive symptoms – Pervasive pattern of negative attributions regarding the partner's intentions, as evidenced by either:
   a. Negative behaviors pervasively attributed to negative personality traits or perceived to be done voluntarily, intentionally, or with negative intent.
   b. Positive behaviors pervasively attributed to temporary states or perceived to be done accidentally, unintentionally, or with hidden negative intentions.
3) Affective symptoms – Interactions with or thoughts about the partner are frequently marked by intense and persistent levels of at least one of the following:
   a. Anger or contempt.
   b. Sadness.
   c. Apathy.

Both criteria A and B are required. Criterion A involves relationship dissatisfaction, comprising three possible presentations (similar to the depressed mood or anhedonia requirement of the major depressive episode criteria): a pervasive and persistent sense of unhappiness with the relationship; persistent thoughts of divorce or separation; a perceived need for professional help for the relationship. Criterion B comprises behavioral, cognitive, and affective symptoms that have appeared repeatedly in the empirical literature; at least one is required.

Table 2 displays the criteria for child physical abuse. Both criterion A (act) and criterion B (impact) are required, as is criterion C (lack of mitigating circumstances – that is, the acts were not committed to protect self from imminent harm, were not part of developmentally appropriate play, and were not committed to protect child from imminent harm).

As in the DSM, where the criteria for "major depressive episode" are separate from, but referenced by, the criteria for "major depressive disorder", the eleven relational disorder criteria have some sub-criteria sets that have proven essential for reliable application of the criteria. As seen in Table 2, "more than inconsequential physical injury" and "more than inconsequential fear reaction" have clarifying criteria. As we discovered during the field trials discussed below, such operationalizations are necessary to achieve high field assessor-master reviewer agreement. Finally, note that criteria B1 (more than inconsequential injury) and B3 (more than inconsequential fear reaction) involve actual impacts, whereas B2 involves potential for more than inconsequential injury. Assessors judge whether the inherent dangerousness of the act, the degree of force used and the physical environment in which the acts occurred constituted a significant potential for serious harm (e.g., pushing a child hard near the top of a flight of stairs, choking an adolescent hard but leaving no bruises).

Table 3 provides the criteria for parenting problems, which follow a similar structure to the other criteria. Criterion A involves substantial parenting difficulties and criterion B involves significant impact on the child from those parenting difficulties. Again, the criteria reflect findings in the empirical literature and provide a basis for inter-rater reliability in clinical settings. Many instances of family dysfunction for which the child now receives a diagnosis (e.g., conduct disorder, oppositional defiant disorder) could also meet criteria for parenting problem.

DEVELOPMENT AND TESTING OF CRITERIA FOR RELATIONAL PROBLEMS

Criteria have been developed for all eleven relational problems. Below we detail the creation and testing of criteria for relational problems related to partner and child maltreatment. The maltreatment criteria were developed in a multistage process described in depth elsewhere (28-30). The steps comprised: a) examining the content validity and field usability of a set of maltreatment criteria already in use; b) creat-

Table 2 Criteria for child physical abuse

A. Non-accidental use of physical force by a child's parent/caregiver. Physical force includes, but is not limited to, spanking with hand; dropping; pushing; shoving; slapping; grabbing or yanking limbs or body; throwing; poking; hair-pulling; scratching; pinching; restraining or squeezing; shaking; biting; throwing objects at; kicking; hitting with fist; hitting with a stick, strap, belt, or other object; scalding; burning; poisoning; stabbing; applying force to throat; strangling or cutting off air supply; holding under water; using a weapon.

B. Significant impact on the child as evidenced by any of the following:
1) More than inconsequential physical injury (see definition below).
2) Reasonable potential for more than inconsequential physical injury (see definition below) given the inherent dangerousness of the act, the degree of force used and the physical environment in which the acts occurred.
3) More than inconsequential fear reaction (see definition below).

C. The acts of physical force were not committed for any of the following reasons:
1) To protect self from imminent physical harm because the child/adolescent was in the act of physical force (see definition below).
2) To play with the child in a developmentally appropriate manner.
3) To protect child or another person from imminent physical harm (including, but not limited to, pushing child out of the way of a car, taking weapon away from suicidal child, stopping child from inflicting injury on another person).

(*Note:* Subsequent actions that were not directly protective – e.g., whipping child for running into the street – would not meet this criterion)

Subcriteria for "More than inconsequential physical injury"

An injury involving any of the following:
A. Any injury to the face or head.
B. Any injury to a child under 2 years of age.
C. More than superficial bruise(s) (i.e., bruise that is other than very light red in color – for example, violet, blue, black – OR bruises with total area exceeding that of the victim's hand OR are tender to light touch).
D. More than superficial cut(s)/scratch(es) (i.e., would require pressure to stop bleeding).
E. Bleeding internally or from mouth or ears.
F. Welt (bump or ridge raised on the skin).
G. Burns.
H. Loss of consciousness.
I. Loss of functioning (including, but not limited to, sprains, broken bones, detached retina, loose or chipped teeth).
J. Heat exhaustion or heat stroke.
K. Damage to internal organs.
L. Disfigurement (including, but not limited to, scarring).
M. Swelling lasting at least 24 hours.
N. Pain felt (a) in the course of normal activities and (b) at least 24 hours after the physical injury was suffered.

Subcriteria for "More than inconsequential fear reaction"

Victim's significant fear reaction, as evidenced by both of the following:
A. Fear (verbalized or displayed) of bodily injury to self or others.
B. At least one of the following signs of fear or anxiety lasting at least 48 hours:
1) Persistent intrusive recollections of the incident.
2) Marked negative reactions to cues related to incident, as evidenced by any of the following:
a. Avoidance of cues.
b. Subjective or overt distress to cues (Note: perpetrator can be a cue).
c. Physiological hyperarousal to cues (Note: perpetrator can be a cue).
3) Acting or feeling as if incident is recurring.
4) Persistent symptoms of increased arousal, as evidenced by any of the following:
a. Difficulty falling or staying asleep.
b. Irritability or outbursts of anger.
c. Difficulty concentrating.
d. Hypervigilance (i.e., acting overly sensitive to sounds and sights in the environment; scanning the environment expecting danger; feeling keyed up and on edge).
e. Exaggerated startle response.

Subcriteria for "Protection of self from imminent physical harm because child was in the act of physical force"

Acts of physical force were committed to protect self from imminent physical harm because the child was in the act of physical force, as evidenced by all three of the following:
A. Act(s) occurred while other was in the act of using physical force. "In the act" begins with the initiation of motoric behavior that typically would result in an act of physical force (for example, charging to hit him/her) and ends when the use of force is no longer imminent.
B. Sole function of act(s) was to stop other's use of physical force.
C. Act(s) used minimally sufficient force to stop other's use of physical force.

ing a unifying conceptualization for what constituted an above-threshold problem; c) reviewing and adapting (where appropriate) existing operationalizations; d) field testing and refining criteria, assessments, and decision-making process; e) testing criteria's use in wide-scale dissemination; f) creating criteria-informed screeners and structured clinical interviews; and g) examining the content validity of the final criteria.

Before describing the results of the multi-stage development/testing process, some context is necessary. First, the criteria were originally developed for use in the US Air Force and have since been adopted across all services of the US Department of Defense and the US Coast Guard. Second, all assessments and diagnostic judgments were conducted with families with maltreatment allegation lodged against them, not with a more general clinical population. Third, the processes used in this context differ slightly from that used in civilian contexts. Although all clinical assessments were conducted by credentialed providers, the decision about whether someone met the criteria was made by a committee.

Step 1: Examine content validity and field usability of existing criteria

Because we were bootstrapping our definitions using existing definitions as a starting point, we conducted two content validity studies using the family maltreatment criteria then in use by the US Department of Defense (31). To maximize the content and clinical validity of the potential criteria, we followed Haynes et al's (32) suggestion to conduct content validity studies using both civilian and military family maltreatment experts (Study 1) and those intended to use the definitions (i.e., field clinicians; Study 2).

Study 1 (28) suggested that the criteria then in force were adequate but could be improved by: a) operationalizing terms, b) eliminating the definitional overlap of emotional abuse and other forms of maltreatment, and c) eliminating the requirement that clinicians predict risk of recurrence to find that incidents met criteria for child emotional abuse or child neglect. In Study 2 (28), field clinicians shared the

Table 3 Criteria for parenting problem

A. Considering the developmental needs of the child, caregiving is markedly outside the bounds of normal, as evidenced by one of the following:
1) Pervasive caregiving difficulties involving either or both of the following:
   a. Underinvolvement (e.g., parent is not bonded to and does not provide loving relationship for the child).
   b. Overinvolvement (e.g., parent is so protective that young adolescent is not afforded any private communication with friends; child is not able to participate in choices about how they will spend their time).
2) Marked difficulties in at least one aspect of parenting, including, but not limited to:
   a. Failure to adequately monitor child (e.g., not supervising a young child's activities; being insufficiently aware of adolescent's activities).
   b. Marked lack of support of, or active interference in, a key major life activity.
   c. Excessive or inappropriate discipline (not meeting criteria for child abuse).
   d. Excessive pressure on child to engage in a single activity or interest (e.g., sport).
   e. Failure to socialize child through nonexistent or poorly enforced limits.

B. Significant impact on the child involving any of the following:
1) More than inconsequential physical injury.
2) Psychological harm, including either:
   a. More than inconsequential fear reaction.
   b. Psychiatric disorder, at or near diagnostic thresholds related to, or exacerbated by, the caregiving difficulty.
3) Stress-related somatic symptoms (related to or exacerbated by the caregiving difficulty) that significantly interfere with child's normal functioning.
4) Reasonable potential for more than inconsequential physical injury due to the inherent dangerousness of the caregiving difficulty and the child's physical environment
5) Reasonable potential for psychological harm. *Note:* The child's level of functioning and the risk and resilience factors present should be taken into consideration.
   a. Reasonable potential for the development of a psychiatric disorder (at or near diagnostic thresholds) due to the caregiving difficulty.
   b. Reasonable potential for significant disruption of the child's physical, psychological, cognitive, or social development due to the caregiving difficulty.

views of the experts that, despite being generally understandable and containing many key elements of maltreatment, the existing definitions were in need of increased operationalization. Further, regarding the process of decision-making, clinicians reported that extra-definitional issues either influenced the decision-making process or caused the decision-making committee to blatantly overrule the definitions.

Step 2: Conceptualize construct and review existing conceptualizations and operationalizations

In developing the maltreatment criteria, Heyman and Slep (28,29) adopted the DSM-like conceptual framework that partner and child physical and emotional abuse and child neglect would require both a specific type of act (e.g., use of physical force for physical abuse) and a significant impact (or high potential for significant impact, such as shooting a gun at a spouse but not hitting him or her). Because of a presumed risk for significant impact of partner sexual abuse and parent-child sexual abuse, the sexual abuse conceptual framework required only a qualifying act.

Step 3: Comprehensively survey and adapt existing legal, research, and clinical definitions and operationalizations

Step 3 involved creating the maltreatment criteria, based in part on the prior US Department of Defense criteria and on operationalizations in existing legal, research, or clinical definitions that were specific enough to promote reliability. Dozens of civilian and military definitions were comprehensively reviewed, including the following: Child Abuse Prevention and Treatment Act (33); Centers for Disease Control (CDC) partner abuse definitions (34); the Modified Maltreatment Classification System (35,36); international and domestic agencies' definitions of child sexual abuse and related terms, including, among many others, those of the NGO Group for the Convention on the Rights of the Child, Focal Point on the Sexual Exploitation of Children (37); the National Incidence Survey on Child Abuse and Neglect definitions (38); state domestic violence statutes (39); state child abuse statutes (40); the US Department of Defense definitions then in force with their proposed modifications (31,41-43); and definitions for partner maltreatment recommended by the Defense Task Force on Domestic Violence (44).

Step 4: Field tests

The next step (Study 3), a pilot field test at five sites (28), aimed to train clinical staff and other case determination committee members at five sites in the use of these definitions; improve iteratively the operationalizations during the field trial; and compare maltreatment decisions with those of master reviewers (i.e., from the Family Translational Research Group at the State University of New York at Stony Brook and from the headquarters of the Air Force Family Advocacy in San Antonio, Texas).

Agreement between committees and the master reviewers was moderate. Based on monitoring meetings prior to the field trial, Heyman and Slep (28) estimated that committee decisions followed the old definitions approximately 50% of the time. Using the new definitions, however, 76% (κ=.48) of allegations were decided by the base committees the same way as they were by master reviewers (Table 4). Although this represented an improvement, neither 76% agreement nor a Cohen's kappa of .48 could be considered adequate.

For the second field trial, Study 4 (28), several changes were made. First, to make the assessment process consistent across sites and assure that the pertinent information was being assessed, we developed a structured clinical interview that paralleled the diagnostic criteria for each form of family maltreatment. The assessing clinicians were provided with and instructed to use these questions. Second, the then-current committees (comprising primarily service providers)

Table 4 Agreement between field decisions and master reviewers

	Pilot field trial			Field trial 2			Dissemination trial		
Type of maltreatment	κ	**Agreement (%)**	**n**	κ	**Agreement (%)**	**n**	κ	**Agreement (%)**	**n**
Partner maltreatment (all types)	.50	74	143	.81	90	320	.85	92	549
Physical	.50	76	103	.82	91	233	.84	92	435
Emotional	.33	70	40	.76	89	79	.83	93	109
Sexual	-	-	0	.75	88	8	.62	80	5
Child maltreatment (all types)	.49	78	184	.87	94	236	.75	88	342
Physical	.55	80	46	.92	96	76	.82	91	115
Emotional	.24	59	27	.89	96	47	.73	90	60
Sexual	.67	83	12	1.00	100	12	.89	95	19
Neglect	.55	81	99	.80	91	101	.66	84	148
Total cases	.48	76	327	.84	92	556	.82	91	891

seemed to have difficulty weighing criterion-pertinent information only; our US Air Force partners decided to change the composition of who served on the determination board. Third, the presentation of information was drastically altered. The former process (used in the pilot field trial) involved a summary by the clinician who completed the assessment, which often did not straightforwardly present criteria-relevant information. In Study 4, all board members who had pertinent information presented it (e.g., assessing clinician, police, work supervisor). Fourth, votes were cast for each criterion separately, aided by a computer-guided decision matrix that presented the diagnostic decisions to be made.

As shown in Table 4, in 92% of cases (κ = .84), decisions in the field about whether or not maltreatment met or exceeded diagnostic thresholds matched those of master reviewers. This exceptional level of reliability suggested that the refinement of the criteria had been successful.

Step 5: Test diagnostic criteria's use in wide-scale dissemination

Although the second trial involved field-generated decisions under real-world conditions, it was clearly an effectiveness trial and did not speak to the performance of the diagnostic system when disseminated broadly. A dissemination trial of the diagnostic system's use under typical conditions at many sites was necessary. Because the diagnostic system was being disseminated worldwide throughout the organization, we were able to randomly select 41 communities to participate in a trial.

As shown in Table 4, agreement between the field-generated decisions and those of master reviewers remained high. The maintenance of adequate agreement is especially noteworthy, given the relative lack of expertise and training among the majority of those using the diagnostic criteria and the standard-operating-procedure nature of the participating sites (i.e., they were not "special volunteer test sites", as in the effectiveness trial). These results are quite encouraging and suggest that diagnostic systems for relational problems can indeed be reliably applied in real world settings despite the discouraging results in the general DSM literature to date (45-48).

Step 6: Select/create criteria-informed screeners and structured clinical interviews

Because clinical adoption of criteria sets is unlikely without measures to aid in screening and structured assessments, we have developed tools for two-stage screening (i.e., quick questionnaire screeners and a set of structured clinical interviews patterned after the Structured Clinical Interview for DSM-IV Axis I Disorders, SCID-I) for use in clinical practice and research for all eleven forms of relational problems listed earlier. In the coming months, we plan to field test these measures in a clinical setting.

Step 7: Examine the content validity of the final criteria

Currently, we are assessing experts' ratings of the content validity of the eleven proposed criteria sets for relational problems. Initial data from the maltreatment data sets indicated that experts had few recommendations for improvements.

CONCLUSIONS AND RECOMMENDATIONS

Nuanced, multifaceted, and content valid diagnostic criteria for relational problems have been created and can be used reliably in the field even by those with little-to-no clinical training/background. The use of definitions such as these would likely lead to more reliable decision making in the field and more consistency across studies. It is notable not just that field decisions reliably agreed with expert decisions for relational problems, but that this agreement was higher than that usually reported for DSM mental disorders. Although not conclusive, the few studies of DSM diagnoses testing the concordance of field clinicians' diagnoses and

"criterion standard" SCID diagnoses have indicated problematic levels of agreement (49). Basco et al (45) reported poor agreement (κ = .13-.45) between DSM diagnoses given by field clinicians and those by master reviewers using structured interviews (sometimes supplemented with other data sources). Kashner et al (46) reported comparable concordance between clinician's diagnoses and SCID diagnoses (κ = .20-.30), except for diagnoses of severe mental illness (κ = .52-.60). Similarly, a study of psychotic first time inpatients (47) reported decent clinician versus SCID/master reviewer agreement (κ = .51-.73), but only for academic and community hospitals; agreement between clinicians at public hospitals and researchers was poor (κ = .13-.34). Contrast this with the field-decision vs. expert reviewer agreement in the maltreatment dissemination trial (κ = .66-.89 across seven forms of maltreatment; κ = .82 overall).

Although our diagnostic criteria were used reliably and demonstrated content validity, further work is necessary to more firmly establish construct validity. We have begun such studies but do not yet have results that can shed light on the validity of the definitions, bringing relational problems in line with the rest of those in the DSM, for better and worse. The former Director of the National Institute of Mental Health, Steven Hyman, summarized nearly universally recognized sentiments about incompletely and variably validated rationally-derived diagnostic criteria: "If a relative strength of DSM is its focus on reliability, a fundamental weakness lies in problems related to validity. Not only persisting, but looming larger, is the question of whether DSM-IV-TR truly carves nature at the joints – that is, whether the entities described in the manual are truly 'natural kinds' and not arbitrary chimeras" (50). One could add relational problems to the list of reliable rationally-derived diagnoses in search of proof that they are natural kinds, and if they are, that the current criteria optimally distinguish them. Future studies must be conducted to establish the convergent, discriminate, discriminative, and predictive validity of the criteria.

Finally, taxometric methods should be used to investigate if making qualitative distinctions is empirically supportable. Some early work indicates that this is true for partner relational problems (51,52) and has been speculated in many quarters for partner maltreatment (53).

In conclusion, clinically significant behavioral or psychological syndromes or patterns occur between family members and are associated with present distress or disability or with a significant increased risk of suffering death, pain, disability, and important losses of freedom. The literature briefly reviewed above incontrovertibly documents both relational problems' syndromes/patterns and their serious sequelae. Criteria for eleven such relational problems – modeled after current DSM diagnoses – have been developed, along with screener questionnaires and SCID-like structured clinical interviews that operationalize the criteria for each problem. The most studied subset of relational problem criteria – those for partner and child maltreatment – have been shown to have remarkably high levels of reliability when used in the field, at agreement levels never reached by DSM diagnoses for individuals.

Science, service, families, individuals, and the DSM itself, would be well served to include diagnostic criteria for relational problems and to consider the various options for placement of relational problems/processes in the DSM-V.

Acknowledgements

The empirical research discussed in this article was supported by US Air Force/US Department of Agriculture contracts CR-19191B-545810 and CR-19191-428142. We also gratefully acknowledge a grant from the Fetzer Institute, which has partially supported the process of writing this article and is also partially supporting the development of the relational problems screeners and structured interview collection/development. The opinions expressed in this article are solely those of the authors and do not necessarily represent the official views of the US Government, the US Department of Defense, or the US Department of the Air Force.

References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th ed., text revised. Washington: American Psychiatric Association, 2000.
2. Widiger TA, Frances AJ, Pincus HA. DSM–IV sourcebook, Vol. 3. Washington: American Psychiatric Association, 1997.
3. Heyman RE, Feldbau-Kohn SR, Ehrensaft MK et al. Can questionnaire reports correctly classify relationship distress and partner physical abuse? J Fam Psychol 2001;15:334-46.
4. Slep AMS, Heyman RE. Where do we go from here? Moving toward an integrated field of family violence. Aggression and Violent Behavior 2001;6:353-6.
5. Heyman RE, Slep AMS. Creating and field-testing diagnostic criteria for partner and child maltreatment. J Fam Psychol 2006;20:397-408.
6. Straus MA. The Conflict Tactics Scales and its critics: an evaluation and new data on validity and reliability. In: Straus MA, Gelles RJ (eds). Physical violence in American families: risk factors and adaptation to violence in 8,145 families. New Brunswick: Transaction Publishing, 1990:3-16.
7. White JW, Smith PH, Koss MP et al. Intimate partner aggression – What have we learned? Comment on Archer. Psychol Bull 2000;126:690-6.
8. Kupfer DJ, First MB, Regier DE. Introduction. In: Kupfer DJ, First MB, Regier DE (eds). A research agenda for DSM–V. Washington: American Psychiatric Association, 2002:xv–xxiii.
9. Mrazek PJ, Haggerty RJ. Illustrative preventive intervention research programs. In: Mrazek PJ, Haggerty RJ (eds). Reducing risks for mental disorders: frontiers for preventive intervention research. Washington: National Academy Press, 1994:215-313.
10. Beach SRH, Wamboldt M, Kaslow NJ et al. Describing relationship problems in DSM-V: toward better guidance for research and clinical practice. J Fam Psychol 2006;20:359-68.
11. Beach SRH, Wamboldt M, Kaslow NJ et al. Relational processes and mental health: a bench to bedside dialogue to guide the DSM-V. In: Beach SRH, Wamboldt M, Kaslow N (eds). Relational processes and DSM-V: neuroscience, assessment, prevention, and treatment. Washington: American Psychiatric Press, 2006:1-20.

12. Whisman MA, Bruce ML. Marital dissatisfaction and incidence of major depressive episode in a community sample. J Abnorm Psychol 1999;108:674-8.
13. Whisman MA. Marital dissatisfaction and psychiatric disorders: results from the National Comorbidity Survey. J Abnorm Psychol 1999;108:701-6.
14. Beach SRH, Kim S, Cercone-Keeney J et al. Physical aggression and depression: gender asymmetry in effects? J Soc Pers Relat 2004; 21:341-60.
15. Cano A, O'Leary KD. Infidelity and separations precipitate major depressive episodes and symptoms of nonspecific depression and anxiety. J Consult Clin Psychol 2000;68:774-81.
16. Kendler KS, Hettema JM, Butera F et al. Life event dimensions of loss, humiliation, entrapment, and danger in the prediction of onsets of major depression and generalized anxiety. Arch Gen Psychiatry 2003;60:789-96.
17. Reiss D, Pedersen NL, Cederblad M et al. Genetic probes of three theories of maternal adjustment: I. Recent evidence and a model. Fam Process 2001;40:247-59.
18. Caspi A, Sugden K, Moffitt TE et al. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. Science 2003;301:386-99.
19. Beach SRH. Marital and family processes in depression: a scientific foundation for clinical practice. Washington: American Psychological Association, 2001.
20. O'Farrell TJ, Fals-Stewart W. Alcohol abuse. J Marital Fam Ther 2003;29:121-46.
21. Miklowitz DJ. The role of family systems in severe and recurrent psychiatric disorders: a developmental psychopathology view. Dev Psychopathol 2004;16:667-88.
22. Repetti RL, Taylor SE, Seeman TE. Risky families: family social environments and the mental and physical health of offspring. Psychol Bull 2002;128:330-66.
23. Cadoret RJ, Winokur G, Langbehn D et al. Depression spectrum disease: I. The role of gene- environment interaction. Am J Psychiatry 1996;153:892-9.
24. Tienari P, Wynne LC, Moring J et al. The Finnish adoptive family study of schizophrenia: implications for family research. Br J Psychiatry 1994;164:20-6.
25. Liu D, Diorio J, Tannenbaum B et al. Maternal care, hippocampal glucocorticoid receptors, and hypothalamic–pituitary–adrenal responses to stress. Science 1997;277:1659-62.
26. Suomi SJ. Developmental trajectories, early experiences, and community consequences: lessons from studies with rhesus monkeys. In: Keating D, Hertzmann C (eds). Developmental health and the wealth of nations: social, biological, and educational dynamics. New York: Guilford, 1999:185-200.
27. Gallo LC, Troxel WM, Matthews KA et al. Marital status and quality in middle-aged women: associations with levels and trajectories of cardiovascular risk factors. Health Psychol 2003;22:453-63.
28. Heyman RE, Slep AMS. Creating and field-testing diagnostic criteria for partner and child maltreatment. J Fam Psychol 2006;20:397-408.
29. Slep AMS, Heyman RE. Creating and field-testing child maltreatment definitions: improving the reliability of substantiation determinations. Child Maltreat 2006;11:217-36.
30. Heyman RE, Slep AMS. Reliability of family maltreatment diagnostic criteria: 41 site dissemination field trial. Submitted for publication.
31. Office of the Under Secretary of Defense. DoD directive 6400.2. Washington: Office of the Under Secretary of Defense for Personnel and Readiness, 1997.
32. Haynes SN, Richard DCS, Kubany ES. Content validity in psychological assessment: a functional approach to concepts and methods. Psychol Assess 1995;7:238-47.
33. US Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau, Office on Child Abuse and Neglect. The child abuse prevention and treatment act. Washington: US Department of Health and Human Services, 2003.
34. Saltzman LE, Fanslow JL, McMahon PM et al. Intimate partner violence surveillance: uniform definitions and recommended data elements, Ver. 1.0. Atlanta: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, 1999.
35. Barnett D, Manly JT, Cicchetti D. Continuing toward an operational definition of psychological maltreatment. Dev Psychopathol 1991; 3:19-29.
36. Barnett D, Manly JT, Cicchetti D. Defining child maltreatment: the interface between policy and research. In: Cicchetti D, Toth SL (eds). Advances in applied developmental psychology: child abuse, child development and social policy. Norwood: Ablex Publishing, 1993:7-73.
37. NGO Group for the Convention on the Rights of the Child, Focal Point on the Sexual Exploitation of Children. Definitions of child sexual exploitation and related terms. NGO Group for the Convention on the Rights of the Child, 2000.
38. Sedlak AJ, Broadhurst DD. Third national incidence study of child abuse and neglect. Washington: National Center on Child Abuse and Neglect, 1996.
39. National Clearinghouse on Child Abuse and Neglect. Compendium of laws, definitions of domestic violence. National Clearinghouse on Child Abuse and Neglect, 2003.
40. National Clearinghouse on Child Abuse and Neglect. Compendium of Laws, Reporting Laws: definitions of Child Abuse and Neglect. National Clearinghouse on Child Abuse and Neglect, 2003.
41. Office of the Under Secretary of Defense. Proposed DoD directive 6400.2 (draft). Unpublished manuscript, May 25, 1999.
42. Office of the Under Secretary of Defense. Proposed DoD directive 6400.2 (draft). Unpublished manuscript, December 25, 2000.
43. Office of the Under Secretary of Defense. Proposed DoD directive 6400.2 (draft). Unpublished manuscript; February 22, 2002.
44. Defense Task Force on Domestic Violence. Initial report of the Defense Task Force on Domestic Violence. Defense Task Force on Domestic Violence, 2001.
45. Basco MR, Bostic JQ, Davies D et al. Methods to improve diagnostic accuracy in a community mental health setting. Am J Psychiatry 2000;157:1599-605.
46. Shear M, Greeno C, Kang J et al. Diagnosis of nonpsychotic patients in community clinics. Am J Psychiatry 2000;157:581-7.
47. Kashner T, Rush A, Suris A et al. Impact of structured clinical interviews on physicians' practices in community mental health settings. Psychiatr Serv 2003;54:712-8.
48. Fennig S, Craig T, Tanenberg-Karant M et al. Comparison of facility and research diagnoses in first-admission psychotic patients. Am J Psychiatry 1994;151:1423-9.
49. First MB. The reliability of psychiatric diagnoses: counterpoint – There isn't enough evidence available to speculate on the reliability of diagnoses in clinical settings. Psychiatry 2007;4:22-5.
50. Hyman SE. Foreward. In: Phillips KA, First MB, Pincus HA (eds): Advancing DSM: dilemmas in psychiatric diagnosis. Washington: American Psychiatric Association, 2003:xi-xxi.
51. Beach SRH, Fincham FD, Amir N et al. The taxometrics of marriage: is marital discord categorical? J Fam Psychol 2005;19:276-85.
52. Whisman MA, Beach SRH, Snyder DK. Is marital discord taxonic and can taxonic status be assessed reliably? Results from a national, representative sample of married couples. J Consult Clin Psychol 2008;76:45-755.
53. Heyman RE, Slep AMS. Relational diagnoses: from reliable rationally-derived criteria to testable taxonic hypotheses. In: Beach SRH, Wamboldt M, Kaslow N et al (eds). Relational processes and DSM-V: neuroscience, assessment, prevention, and treatment. Washington: American Psychiatric Press, 2006:139-56.